Detecting Medical Simulation Errors with Machine learning and Multimodal Data

Detecting Medical Simulation Errors with Machine learning and Multimodal Data

Daniele Di Mitri

Welten Institute, Research Centre for Learning, Teaching and Technology
Open University of the Netherlands
Daniele.Dimitri@ou.nl

Abstract. In this doctoral consortium paper, we introduce the CPR Tutor, an intelligent tutoring system for cardiopulmonary resuscitation (CPR) training based on the analysis of multimodal data. Using a multi-sensor setup, the CPR Tutor tracks the CPR execution of the trainee and generates automatic adaptive feedback to improve the trainee's performance. This research work is part of a PhD project entitled "Multimodal Tutor: adaptive feedback from multimodal experience capturing", a project which investigates how to use multimodal and multi-sensor data to generate personalised feedback for training psycho-motor skills at the workplace or during medical simulations. In the CPR Tutor, we use Microsoft Kinect and Myo to track trainee's body position and the ResusciAnne QCPR manikin to get correct CPR performance metrics. We then use a validated approach, the Multimodal Pipeline, for the collection, storage, processing, annotation of multimodal data. This paper describes the preliminary results obtained in the first design of the CPR Tutor.

1 Background

Learning how to master practical skills is highly relevant in the medical domain. Often healthcare professionals are not given enough opportunities for training their medical skills. Existing simulations quite often lack on-task, real-time and actionable feedback, which results in longer and less effective training for the professionals. With the Multimodal Tutor, we aim at supporting skills training and simulations in a variety of learning scenarios using multimodal data, machine learning and intelligent tutoring. With the term multimodal data, we refer to 1) micro-level learners behavioural data: i.e. motoric actions or physiological responses; 2) data of the learning situation such as learning context, environment and activity. Such data can be collected using wearable sensors, cameras or Internet of Things (IoT) devices. Combining data from multiple modalities allows us to achieve a more accurate representation of the learning process. If multimodal data are associated with performance measurements, they can be fed into machine learning algorithms which can detect training errors and which can be used to provide on-time automatic and personalised feedback. In our PhD project, we have looked into two relevant medical use-cases, among the several medical scenarios where to use the Multimodal Tutor.

The first use-case, developed within the WEKIT Project[1] consists of learning how to perform different eco-graphic tests using an Ultrasound Machine for acquiring medical images of, for example, the carotid or spleen test. The second example, investigated in depth in this research work, deals with cardiopulmonary resuscitation training, more precisely how to correctly perform chest compression (CCs) as part of the CPR procedure. We selected CPR as a representative task for multiple reasons: it is usually trained singularly and has a clearly defined procedure. The learning objectives are also well defined: e.g. the compressions should be around 120 per minute and 5 cm deep. Finally, CPR has a strong relevance: the cases of cardiopulmonary arrest are unfortunately very common. The more people are trained to do CPR the higher will be the chance of saving lives.

2 Related studies

Examples of multimodal tutors, i.e. intelligent tutors using multimodal data to provide adaptive and personalised feedback are AutoTutor [5] or the Affective Learning Companion [1]. For the specific case of Multimodal Tutors for CPR, previous researchers have used Kinect based systems for tracking the CPR performance. [7] first piloted a Kinect-based system for providing feedback. The study in [8] designed a Kinect-based real-time audiovisual feedback device to investigate the relationship between rescuer posture, body weight, and CC quality. They tested 100 participants monitoring depth and rate of CC and providing further real-time feedback. The result of this study was that kneeling posture provides better CC than a standing posture and that audio-visual feedback can provide a better CC depth, rate, and effective CC ratio. In our study we proposed the use of neural network in order to detect training errors in terms of CC rate, CC depth, CC release, but also to detect additional training errors, so far not tracked by commercial manikins, such as the correct locking of the arms during and the correct use of body posture and body weight during CC.

3 Goal and Research Questions

The goal of the research work presented in the paper is to further develop the idea of a CPR Tutor [2], a specific implementation of the Multimodal Tutor, and answer the following research questions:

1. Is possible to classify CPR training errors using multimodal sensor data?
2. Can we use multimodal data to classify custom defined CPR errors?
3. Can we train classifiers generalisable across different users and setups?

[1] Wearable-enhanced Knowledge Intensive Training - WEKIT project http://www.wekit.eu

Fig. 1. The graphic representation of the experimental setting

4 Proposed approach

To set the theoretical ground of the Multimodal Tutor, we have first proposed a conceptual model, the *Multimodal Learning Analytics Model* (MLeAM)[4]. MLeAM shows how multimodal data can produce feedback to the learner and it helps to classify similar experiments in the field. On the technical side, we developed a technological system, the *Multimodal Pipeline*, which consists in a chain of multiple technological solutions for the collection, storing, processing, annotation and exploitation of multimodal data for supporting learning.

The first component of the Pipeline is the *Multimodal Learning Hub* [6] an Open Source software which collects data from multiple sensor applications, synchronises them and store them into a custom data format named *Meaningful Learning Task* (MLT). For the case of the CPR Tutor, the Learning Hub collects 1) body position from Microsoft Kinect; 2) electro myogram information from the Myo armband.

The second component is the *Visual Inspection Tool* [3] a web application which allows to visualize and annotate the MLT sessions collected with the LearningHub. For the case of the CPR Tutor, the VIT allows plotting the Kinect and Myo along with the video recordings. It also allows loading the data from the ResusciAnne manikin's wireless Simpad SkillReporter which are used as baseline annotations for the CPR Tutor. The Simpad SkillReporter records start and stop time for each single CC and rate them according to the their 1) *Compression Rate* (ideal values: 100-120 per minute); 2) *Compression Depth* (ideal values: 5-6 cm); 3) *Compression Release* (ideal values: 0-1 cm). An idea how the compressions are annotated in the VIT is shown in figure 2. Moreover, besides loading the performance measurements which can be retrieved by the Simpad SkillReporter, the VIT allows defining expert-driven custom annotations. For example, each CC can be marked with additional training errors that the ResusciAnne manikin is not able to track such as if the arms are locked or the body weight is used correctly.

Once the MLT session has been annotated it can be loaded into the third component of the Multimodal Pipeline, a data processing script that processes the annotated dataset and prepares it for supervised machine learning. The script can process multiple sessions at the time.

5 Results

In December 2018, we ran our first experimental trial for collecting the necessary data to design the CPR Tutor. We tested 11 experts (medical students) performing one at the time two sessions of two minutes CCs. We collected in total around 5200+ CC. The setup of the experiment is similar to the one represented in figure 1. The dataset obtained includes 53 attributes (time series) such as *ElbowRightX*, *ShoulderLeftZ* as each body jointure for Kinect is a 3D vector.

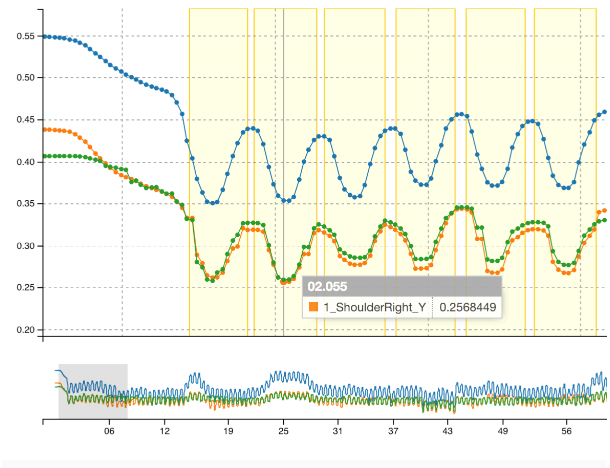

Fig. 2. Plotting example of the CPR Tutor data. The yellow columns are time-intervals which indicate the start and end of the chest-copressions used as training samples. Each chest-compression is rated with three classes.

All CCs were labelled for CompressionRate and CompressionDepth with three classes: 0 - too slow, 1 - correct, 2 - too fast and CompressionRelease with two classes 0 - wrong 1 - right. These classes are based on the type of feedback meaningful for the learner e.g. either ‘too slow’, ‘on-point’ or ‘too fast’. These perfomance metrics were extracted by the ResusciAnne manikins. To prepare the data for machine learning we considered each of the 5254 CCs as one learning sample. Each of these samples is a matrix of 53 attributes times 8, which is a fixed number of bins by which the CCs were re-sampled to make them even in length. That lead having a tensor of shape (5254, 8, 53) And corresponding 1D output vector of labels of size 5254.

We fed this tensor into three Recurrent Neural Networks using the TensorFlow-Keras library. We split the dataset into 2/3 training and 1/3 test. We trained three different classifiers as we have three target classes. In the networks we used a sequential model having first a Long Short Term Memory network of shape (8,53) and then a dense layer which with a SoftMax function classify the most likely class among the 3 or 2 options. With such setup, we were able to get reasonable accuracy of the model of around 70-75%.

6 Questions for the Doctoral Consortium

In this paper we presented our approach for the design of the CPR Tutor. In the current approach we used each CC as training sample by masking/windowing of the original time series. Then we trained an LSTM network with all these samples. With this approach we are able to classify accurately the target classes, however discarding the rest of the time-series we are not able to detect if a CC happened. We ask the DC how, given the available data, could we train a classifier able to detect whether a CC happened or not. We also ask the DC to provide feedback on the methodology and the relevance of the project.

References

1. Burleson, W.: Affective learning companions. Educational Technology 47(1), 28 (2007)
2. Di Mitri, D.: Multimodal tutor for CPR. In: Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). vol. 10948 LNAI, pp. 513–516. Springer International Publishing, Cham, Switzerland (2018)
3. Di Mitri, D., Schneider, J., Klemke, R., Specht, M., Drachsler, H.: Read Between the Lines: An Annotation Tool for Multimodal Data for Learning. In: Proceedings of the 9th International Conference on Learning Analytics & Knowledge - LAK19. pp. 51–60. ACM, New York, NY, USA (2019), http://dl.acm.org/citation.cfm?doid=3303772.3303776
4. Di Mitri, D., Schneider, J., Specht, M., Drachsler, H.: From signals to knowledge: A conceptual model for multimodal learning analytics. Journal of Computer Assisted Learning 34(4), 338–349 (aug 2018), http://doi.wiley.com/10.1111/jcal.12288
5. D’Mello, S., Jackson, T., Craig, S., Morgan, B., Chipman, P., White, H., Person, N., Kort, B., El Kaliouby, R., Picard, R.W., Graesser, A.: AutoTutor detects and responds to learners affective and cognitive states. IEEE Transactions on Education 48(4), 612–618 (2008)

6. Schneider, J., Di Mitri, D., Limbu, B., Drachsler, H.: Multimodal Learning Hub: A Tool for Capturing Customizable Multimodal Learning Experiences. In: Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). vol. 11082 LNCS, pp. 45–58. Springer, Cham, Switzerland (2018)
7. Semeraro, F., Frisoli, A., Loconsole, C., Bannò, F., Tammaro, G., Imbriaco, G., Marchetti, L., Cerchiari, E.L.: A new Kinect-based system for the analysis of performances in cardiopulmonary resuscitation (CPR) training. Resuscitation 83, e20 (2012)
8. Wang, J.C.C., Tsai, S.H.H., Chen, Y.L., Chu, S.J.J., Liao, W.I.I.: Kinect-based real-time audiovisual feedback device improves CPR quality of lower-body-weight rescuers. American Journal of Emergency Medicine 36(4), 577–582 (apr 2018), http://www.ncbi.nlm.nih.gov/pubmed/28927950 https://linkinghub.elsevier.com/retrieve/pii/S0735675717307477 https://doi.org/10.1016/j.ajem.2017.09.022